DEPARTMENT OF HEALTH AND HUMAN SERVICES PUBLIC HEALTH SERVICE

FOOD AND DRUG ADMINISTRATION
CENTER FOR BIOLOGICS EVALUATION AND RESEARCH

Division of Biostatistics and Epidemiology (HFM-215)

Memorandum

DATE: June 30, 1997

FROM: *Teresa Neeman, Ph.D.* Teresa Neeman

THROUGH: *Peter A. Lachenbruch, Ph.D., Chief, Biostatistics Branch* Peter Lachenbruch

SUBJECT: *Biologic License Application 96-1433,*
Recombinant Human Interleukin Eleven (rhIL-11) for the Reduction of Severe Thrombocytopenia in Subjects undergoing Mylosuppressive Chemotherapy,
Genetics Institute

TO: *Richard Steffen, MD*
Clinical Reviewer, Division of Clinical Trial Design and Analysis (DCTDA)

INTRODUCTION

The statistical review for BLA 96-1433 covers three double-blind placebo controlled clinical trials designed to evaluate the safety and efficacy of rhIL-11 in patients with chemotherapy induced thrombocytopenia. Two of the trials, 9308 and 9416, were submitted as supportive of efficacy. The third trial, 9313, was submitted as part of the safety database. Each of the trials was designed to support a different indication. The first trial, 9308, was designed to evaluate the claim that rhIL-11 reduces the need for platelet transfusion in patients who had required a transfusion in a previous cycle. An enrolled patient received the same type and dose of chemotherapy given in the previous cycle. The entry criteria were fairly broad, and as a results many types and doses of chemotherapy were represented. In this study, the proportion of patients who avoided platelet transfusions were compared between the arms. The second trial, 9313, was designed to evaluate the efficacy of rhIL-11 in a bone marrow transplant setting. Because it was anticipated that all patients would be transfused, the primary endpoint was time to platelet recovery. The third trial, 9416, enrolled patients with breast to receive study drug over two cycles of chemotherapy. The primary endpoint was originally the number of platelet transfusions. After

one-third of the patients had been enrolled, a blinded analysis of the primary endpoint indicated that the actual transfusion rate was much less than anticipated. In consultation with the FDA, the primary endpoint was changed to the need for platelet transfusion, defined as a platelet count at or below 20,000/μL.

SUMMARY OF CONCLUSIONS

The results of these studies support the activity of rhIL-11 in reducing the need for platelet transfusions in a restricted setting. For myeloablative chemotherapy seen in the bone marrow transplant study, patients randomized to placebo were indistinguishable from patients randomized to rhIL-11 with respect to the primary endpoint. In trial 9308, a treatment effect was seen among patients receiving less myelosuppressive chemotherapies. There was scant evidence of a treatment effect in patients receiving more myelosuppressive chemotherapies. In the trial 9416, fewer than half the patients were transfused over both cycles. Differences were seen in the primary endpoint between the placebo arm and the rhIL-11 arm which approached statistical significance. Although this trial may not stand alone as a single confirmation of treatment effect, it is supportive of the results seen in trial 9308.

STUDY 9308

I. BACKGROUND

Study 9308 was a multi-center double-blind placebo-controlled three-arm trial designed to evaluate the efficacy of rhIL-11 in secondary prophylaxis for severe thrombocytopenia in subjects undergoing mylosuppressive chemotherapy. The 93 enrolled patients who had severe thrombocytopenia in their prior cycle of chemotherapy were randomized to one of three arms: placebo, low dose rhIL-11 (25 μg/kg) or high dose rhIL-11 (50 μg/kg). According to the protocol, each patient was to receive the same dose of chemotherapy they had received in the previous cycle, followed by study drug or placebo, and followed for at least 30 days. The randomization was stratified by the number of days of the chemotherapy cycle (Longer vs. Shorter) and prior chemotherapy (More vs. Less). After the masked cycle, patients could go on an open-label extension of the study and receive drug following additional chemotherapy cycles. The primary endpoint in this study was the incidence of platelet transfusions in the masked cycle, which were given when platelet counts were ≤20,000/μL.

The study population was heterogenous. The types of malignancies represented included (but not limited to) breast cancer, Non-Hodgkin's Lymphoma, ovarian cancer, small cell lung cancer, Hodgkin's disease and non-small cell lung cancer. The type and dose of chemotherapy the patient received on study was the same type and dose chemotherapy they had received in the previous cycle. There were 24 different types of chemotherapy regimens represented in this study, and even under the same regimen, doses may have varied from patient to patient. Some chemotherapies, however, such as Dose-intense cyclophosphamide, etoposide and cisplatin (DiCEP) which were all administered by the same investigator, probably used a uniform dose.

II. PATIENT DISPOSITION

Patients Not Meeting Inclusion Criteria

Of the 93 patients enrolled, 22 patients did not meet inclusion criteria. Only 1 patient, patient # 93, was considered not to be evaluable based on ineligibility. This patient had a platelet count at baseline below 100,000/μL. It is not clear to this reviewer why this patient was selectively excluded from the evaluable patients, but others were included. The list of patient numbers and reasons for not meeting the inclusion criteria are displayed in the table below:

Table 1: Patients not Meeting Inclusion Criteria, Study 9308

Inclusion Criterion not met	Patient #
Patient did not receive platelet transfusion following previous chemotherapy cycle	1, 83
Patient did not have adequate renal/hepatic function (as specified in Case Report Form, Inclusion Criteria, volume 48, p. 281)	10, 12, 15, 45, 67, 71, 88, 90
Patient did not have performance status of 0 or 1	32, 44, 73
Baseline Hemoglobin < 9.5 g/dL	42, 76, 79, 82, 83
Platelet Count at baseline < 100,000/μL	42, 93
ANC at baseline < 1000/mm^3	50, 54
Patient less than 18 years old	70

Of the 93 enrolled patients, 5 withdrew consent prior to receiving either study drug or placebo (patient #s 4, 15, 21 in the high dose arm and # 8,12 in the low dose arm). These patients were not considered to be evaluable in the evaluable patient population. Six other patients were not considered to be evaluable because of major protocol violations or failing to meet eligibility criteria. The table below summarizes the reasons for not including subject in the evaluable patient population.

Table 2: Patients not considered Evaluable for Primary Analysis, Study 9308

	Patient #	reason patient unevaluable
Placebo	30	chemotherapy reduced
	54	received platelet transfusion with platelet count >20,000/μL
	93	baseline platelet count < 100,000/μL
low dose rhIL-11	89	chemotherapy reduced
high dose rhIL-11	3	patient not transfused with platelet count <20,000/μL
	83	received platelet transfusion with platelet count >20,000/μL

III. SPONSOR'S ANALYSIS

The sponsor based the primary endpoint analysis on three populations. The principal analysis was the evaluable patient population, consisting of the 82 patients as defined in the section above: 27 in the placebo arm, 28 in the low dose arm, and 27 in the high dose arm. Supporting analyses were done on the intent-to-treat population, consisting of all 93 enrolled patients (30 in the placebo arm, 31 in the low dose arm, and 32 in the high dose arm), and on the "completers" population consisting of the 69 patients who completed study drug administration (25 in the placebo arm, 25 in the low dose arm, and 19 in the high dose arm).

This reviewer verified from the electronic line listings that, in the evaluable patient population, that the number of patients in each arm listed as having received a platelet transfusion agreed with the summary statistics in the study report. It was also confirmed that 5 patients never received study drug, and six (6) had the protocol violation described above.

The protocol (Volume 48, Appendix A) states that the statistical analysis of the primary endpoint will be based upon the evaluable patient population, defined as all enrolled patients, except "those who discontinue the study before receiving a full course of rhIL-11 due to either cancer progression or due to toxicity unrelated to rhIL-11; or those for which a major protocol violation occurred during treatment." The protocol goes on to state that if any of these patients requires a platelet transfusion, he will be considered a treatment failure in the primary analysis. Because this was intended as a Phase II trial, the sponsor did not state prospectively how patients not evaluable for the primary endpoint would be treated in an Intent-to-Treat analysis. It was stated that prior to unmasking, rules were developed for how these patients would be evaluated. If there was no record of a platelet transfusion, and few or no daily platelet counts, but any existing counts were above 20,000/μL, the patient was considered to be a treatment success. This method of assigning response in the face of little or no data worked in favor of a treatment effect, as one can see from the table below.

The statistical analysis proposed by the sponsor consisted of comparing placebo with each dose of drug using a two-sided Fisher's Exact Test. Since two comparisons were to be made, the sponsor subsequently adjusted the p-values for multiplicity by bootstrap resampling using PROC MULTTEST in SAS. A summary of the sponsor's analyses for the primary endpoint appears in the table below.

Table 3: Sponsor's Analyses of Primary Endpoint, Overall Incidence of Transfusion, Study 9308

	Placebo Arm (successes/total)	**low dose rhIL-11 (successes/total)**	**high dose rhIL-11 (successes/total)**	**p-value high vs. placebo low vs. placebo (Fisher's Exact Test)**
Evaluable Patient Population	1/27	5/28	8/27	Adjusted .02 .02 .19 .22
Intent-to-Treat Population	2/30	6/31	12/32	.005 .006 .26 .34
Completing Patient Population	0/25	4/25	7/19	.001 .001 .11 .10

IV. Reviewer's Analysis of Primary Endpoint

A. Comments on Completers Analysis: This reviewer can understand the rationale for presenting an evaluable patient analysis and an Intent-to-Treat analysis. When there are missing data, one usually does not know if their missingness is related to what their response might have been if it had been observed. If the reason the patient dropped out was not related to their "missing" response and not related to a treatment effect, then inference based upon the patient with a known response (the evaluable patient population) is valid and the estimate of the underlying treatment effect is unbiased. If patients are dropping out because of drug toxicities or inconveniences related to the drug, then an evaluable patient analysis may be biased in favor of the experimental treatment. For this reason, it is appropriate to support this analysis with an Intent-to-Treat analysis which should treat patients with unknown outcome in a conservative manner so as to diminish treatment effect. If a treatment effect persists in the face of this more conservative analysis, one feels more confident that the treatment effect seen despite the missing data was real.

The rationale for a completers analysis is to measure the treatment effect when the drug is

administered as intended. It is assumed that the patients who dropped out do not contribute meaningful information or contribute biased information to the endpoint Analyzing the data using only the completer's information provides an estimate of what the effect size of the drug on the endpoint is under optimal conditions. However, for drug approval, the interest is on the drug's effect on the patient which is only incompletely measured by the endpoint. If a patient drops out because of drug-related toxicities or drug intolerance, this information should be reflected in the measure of drug efficacy. Moreover, the "completers" population defined in this submission were patients who did not complete study dosing. If one of these patients had a transfusion while still on drug, then the patient should have been counted as a treatment failure even by a completer's analysis, because the impact of the drug on the endpoint can still be measured. This reviewer, therefore, will not consider a "completer's" analysis to be an appropriate measure of overall treatment effect.

B. Comments on the Evaluable Patient Analysis: The primary efficacy analysis was based upon 82 out of 93 patients deemed by the sponsor to be evaluable. The analysis was based upon overall incidence of transfusion in each of the 3 arms. This reviewer noted that randomization was not stratified by site, and consequently there were imbalances between treatment groups within centers. In 2 of the 20 centers, there were no evaluable patients. Five of the centers had only one evaluable patient. Two of the centers, with 2 and 3 evaluable patients respectively, treated all patients with high dose rhIL-11. Of the remaining 11 centers, there were 5 centers in which either the high dose arm or the placebo arm was not represented, so that a comparison between high dose and placebo could not be made. A breakdown of the incidence of platelet transfusion is given in the table below:

Table 4: Incidence of Transfusion by Site, Study 9308

	transfused	avoided transfusion	not evaluable
Site # 33 (N=19)			
Placebo	7	0	0
low dose rhIL-11	6	0	0
high dose rhIL-11	6	0	0
Site # 23 (N=19)			
Placebo	8	0	1
low dose rhIL-11	6	0	0
high dose rhIL-11	2	1	1
Site #32 (N=7)			
Placebo	1	0	0
low dose rhIL-11	0	3	0
high dose rhIL-11	1	2	0

	transfused	avoided transfusion	not evaluable
Site #31 (N=6)			
Placebo	2	0	0
low dose rhIL-11	1	0	0
high dose rhIL-11	3	0	0
Site #56 (N=6)			
Placebo	3	0	1
low dose rhIL-11	2	0	0
high dose rhIL-11	0	0	0
Site #30(N=5)			
Placebo	1	0	0
low dose rhIL-11	2	0	0
high dose rhIL-11	1	0	1
Site #63 (N=4)			
Placebo	1	1	1
low dose rhIL-11	0	1	0
high dose rhIL-11	0	0	0
Site #39 (N=4)			
Placebo	0	0	0
low dose rhIL-11	0	0	0
high dose rhIL-11	1	2	1
Site #25 (N=4)			
Placebo	3	0	0
low dose rhIL-11	0	0	0
high dose rhIL-11	1	0	0
Site #26 (N=3)			
Placebo	0	0	0
low dose rhIL-11	1	0	1
high dose rhIL-11	0	1	0
Site #38 (N=3)			
Placebo	0	0	0
low dose rhIL-11	1	0	0
high dose rhIL-11	2	0	0

	transfused	avoided transfusion	not evaluable
Site #7 (N=2)			
Placebo	0	0	0
low dose rhIL-11	1	0	0
high dose rhIL-11	0	0	1
Site # 9 (N=2)			
Placebo	0	0	0
low dose rhIL-11	0	0	0
high dose rhIL-11	1	1	0
Site # 27 (N=2)			
Placebo	0	0	0
low dose rhIL-11	0	1	0
high dose rhIL-11	0	1	0
Site #100 (N=2)			
Placebo	**0**	**0**	**0**
low dose rhIL-11	**0**	**0**	**1**
high dose rhIL-11	**1**	**0**	**0**

In addition, site #28 had 1 patient randomized to high dose rhIL-11, but the patient never received study drug (patient 4). Sites #s 24, 62, 75 and 90 had 1 patient each assigned to receive low dose rhIL-11. Three of these patients were transfused, and one (patient 12) never received study drug.

One notices immediately that the incidence of transfusions varies markedly between centers. In the two largest centers, the transfusion rates among the evaluable patients are 100% and 94% respectively, whereas in the next largest center (N=7), only 29% of the patients were transfused. Of the patients in that center who avoided transfusion, 3 were in the low-dose rhIL-11 arm and accounted for 60% (3/5) of the total number of patients in that arm who avoided transfusion.

However, patients are being measured with respect to their own baseline. Per protocol, nothing about the patients' care should have changed between cycle X and cycle X+1 except the addition of study drug. Differences in transfusion rates among centers may be less a reflection of standard of care at each center (since each of these patients were transfused in the previous cycle) and more a reflection of the mylosuppresiveness of the chemotherapy used at that center and a drug effect in the less mylosuppressive regimens. For example, the 19 patients in center #33 all received DiCEP, a very mylosuppressive regimen, and all patients were transfused. All of these patients had severe neutropenia, and the median time to ANC recovery was 13 days. In contrast, of the seven patients at site # 32, only three (3) had severe neutropenia.

As an exploratory analysis, we investigated the possibility that, while in very mylosuppressive chemotherapy regimens, rhIL-11 may not be effective or only minimally effective, in other less mylosuppressive regimens a significant treatment effect may be seen. Degree of mylosuppression was measured not with respect to the strata defined by the sponsor (Longer vs. Shorter), but rather the amount of mylosuppression each patient experienced as determined by absolute neutrophil counts (ANC). Ideally, we would have measured mylosuppression as days to neutrophil recovery in cycle X. Because we did not have this information, we used the days to neutrophil recovery in cycle X+1. A logistic regression was used in S-PLUS to model the probability of receiving a platelet transfusion using time to ANC>500 as a covariate. The analysis suggested that time to ANC>500 was an important covariate ($p<.01$), where long neutrophil recovery times correspond to a greater need for transfusion. After adjusting for this covariate, another logistic regression analysis was performed, confirming that the treatment effect seen when comparing the high dose rhIL-11 with placebo remains.

In order to make this relation more explicit, evaluable patients were divided into two approximately equal groups according to the number of days to ANC recovery: less than 11 days and greater than or equal to 11 days. The data are summarized in the tables below:

Table 5: Incidence of Transfusion in Patients with ANC recovery <11 Days, Study 9308

	transfused	**avoided transfusion**	**total**
Placebo	1 (8%)	11 (92%)	12
low dose rhIL-11	4 (31%)	9 (69%)	13
high dose rhIL-11	6 (50%)	6 (50%)	12
total	11	26	37

Table 6: Incidence of Transfusion in Patients with ANC recovery ≥11 Days, Study 9308

	transfused	**avoided transfusion**	**total**
Placebo	0 (0%)	15 (100%)	15
low dose rhIL-11	1 (7%)	13 (93%)	14
high dose rhIL-11	2 (13%)	13 (87%)	15
total	3	41	44

C. Comments on Intent-to-Treat Population: The purpose of an intent-to-treat analysis is to re-assess a possible treatment effect in a more conservative evaluation. If patients drop out a study because of drug toxicities or inconveniences associated with the drug, one should include

these patients in order to capture this information.

There were five patients who never received study drug. Two had been randomized to low dose rhIL-11 and the other three, to high dose rhIL-11. The two patients in the low dose arm were transfused; however, there is no documentation for transfusions in the high dose arm. There are almost no platelet count data for any of these patients, so it cannot be determined if a platelet transfusion was ever required. There are two issues of concern. The first is, that in the absence of transfusion and platelet documentation, it should not be assumed that no transfusions were done or required. The second is that it seems unlikely that, of the 5 patients who received no drug, only 2 were transfused, while in the placebo arm 26 out of 27 evaluable patients were transfused. These concerns should not be interpreted as doubting the integrity of the sponsor's interpretation of the data. Rather, this reviewer believes that a treatment effect observed in a more conservative analysis carries more weight than a treatment effect seen in which the treatment arm is always given the benefit of the doubt.

There are two ways of approaching this intent-to-treat analysis. If all randomized patients are included and patients with incomplete data are counted as failures, then the success rates are 2/30, 6/31 and 8/32 for the placebo, low dose and high dose arms respectively. A comparison of the placebo vs. The high dose yields an unadjusted two-sided p-value of 0.08. On the other hand, if only patients who received study drug are included in the intent-to treat analysis, then the success rates are 2/30, 6/29, and 8/29 for the placebo, low dose and high dose arms respectively. The associated (unadjusted) p-value comparing high dose to placebo in a two-sided Fisher's Exact Test is 0.04. Both analyses are supportive of the primary analysis based upon the 82 evaluable subjects.

D. Comments on Adjusting p-values for Multiplicity

P-values are often adjusted when two or more hypotheses are being considered in the same study, the rationale being, that when both hypotheses H_{01} and H_{02} are true, the likelihood of at least one the p-values is less than .05 is greater than the probability that the p-value associated with H_{01} is less than .05. Suppose that the two p-values p_1, p_2 are such that $p_1<p_2$. Then the true significance level of H_{01} under the complete null hypothesis is:

$$P_{H_{01} \cap H_{02}}(\min(P_1,P_2) < p_1)$$

Under Bonferroni, under which no distributional assumptions about min(P_1, P_2) are made, this probability is bounded by $2p_1$. However, under the resampling schemes of Westfall and Young[1], the known distribution of min(P_1, P_2) under the complete null hypothesis is used to give stricter bounds than Bonferroni.

---

[1]P. Westfall & S. Young, Resampling-Based Multiple Testing, J. Wiley and Sons, 1993.

In this study, p-values were adjusted by the sponsor using a bootstrap resampling method in PROC MULTTEST. The adjusted p-value for the comparison of the high dose rhIL-11 and placebo was 0.023, although the raw p-value was 0.024.

At first glance, this appears to be an impossible conclusion; as was stated in the first paragraph, any p-value adjustment should be in the upward direction to take into account that multiple testing can increase the likelihood of incorrectly rejecting a null hypothesis. However, it must be recognized that 0.023 is an upward adjustment from the p-value associated with H_{01} when both H_{01} and H_{02} are true. We elaborate in the following paragraph.

The strength of evidence supporting the alternative to H_{01}, when both H_{01} and H_{02} are true, may be different than the evidence in favor of the alternative to H_{01} when only H_{01} is true. The former, expressed as a p-value, is the probability

$$P_{H_{01} \cap H_{02}}(P_1 < p_1)$$

In this study, this p-value can be computed as shown below:

$$\sum_{(a,b)\in S} \frac{\binom{14}{a\ b}\binom{68}{27-a\ 27-b}}{\binom{82}{27\ 27}} = 0.013$$

where the letters a and b are the possible numbers of successful outcomes of placebo and high dose in this experiment, and S is the subset of outcomes which are *at least as extreme* as the outcome of the study. A *more extreme outcome* is one with an associated p-value of no more than 0.024 in a one-to-one comparison of high-dose against placebo (using, for example, a two-sided Fisher's Exact Test).

On the other hand, the strength of the evidence of H_{01} independent of H_{02} is expressed by the probability

$$P_{H_{01}}(P_1 < p_1)$$

which, in this study, can be computed by the formula:

$$\sum_{a \in S} \frac{\binom{9}{a}\binom{45}{27-a}}{\binom{54}{27}} = 0.024$$

where S is the set of all outcomes which are at least as extreme as the outcome of the study.

The adjusted p-value (0.023) from bootstrap resampling is the adjustment of p-value 0.013. The appropriate p-value to report, however is maximum of (0.024, 0.023) because it reflects the strength of the evidence of a treatment effect (high dose vs. Placebo) for the more conservative null hypothesis.

V. Secondary Analyses: RBC transfusions, Platelet transfusions and Total # of Transfusions

These summaries were based upon the 87 patients: 88 patients received study drug, but patient #3 (high dose) had incomplete follow-up and it could not be determined how many platelet transfusion would have been required. It was felt that patients who did not receive study drug did not have enough follow-up to determine transfusion requirements. The sponsor also presented a summary of these data based upon the 82 evaluable patients, which differed slightly from the FDA summary.

The median numbers of platelet transfusions administered were 2.5, 2 and 1 in the placebo, low dose and high dose arms respectively. The mean numbers of platelet transfusions were 3.3, 2.0, and 2.2 . The difference in the numbers of platelet transfusions between the placebo and high dose arm was not found to be statistically significant from either the Wilcoxon Test (p=0.07) or the t-test (p= 0.15).

If a transfusion was not given per protocol, it should have been documented and included in the electronic line listings called "Transfusions". The bioresearch monitor noticed that on several occasions, missed transfusions were not documented. When the platelet data were subsequently reviewed, it was not always clear as to what constituted a missed transfusion. In this review, we accepted the company's documentation as reliable. When "missed" platelet transfusions are added to actual platelet transfusions, the medians were unchanged in each of the arms. The mean numbers of platelet transfusions were 3.6, 2.5, and 2.9 in the placebo, low dose and high dose arms respectively. There were no statistically significant differences seen using either the Wilcoxon Test (p= 0.15) or the t-test (p= 0.44).

The median number of red blood cell transfusions was two (2) in each of the three arms, and the 25%-75% quartile was [1,3] in each of the arms. The mean numbers of red blood cell transfusions were 2.0, 2.2 and 2.0 in the placebo, low dose rhIL-11 and high dose rhIL-11, respectively. These summary statistics came from data extracted from the electronic line listings.

In contrast to the number of red blood cell transfusions received, one can also measure the number of units each patient received. This information came from the SAS database from study 9308. The median numbers of units of red blood cells transfused were 2 (Q1-Q3: 2-4 units) in the placebo arm, 4 (Q1-Q3: 2-4 units) in the low dose arm, and 2 (Q1-Q3: 2-4 units) in the high dose arm. The mean numbers of units transfused were 3.0 (standard deviation 2.8), 3.2 (standard deviation 1.7), and 3.1 (standard deviation 2.0) in the placebo, low dose and high dose arms respectively.

The total number of transfusions of any kind was also compared among the treatment arms. The histograms (Figure 1) indicate that a larger proportion of patients in the high dose rhIL-11 arm experience no more than 2 total transfusions, although the median number of transfusions is similar in the placebo and high dose groups. The differences between the total numbers of transfusions in the high dose group compared with placebo is not statistically significant (Wilcoxon Rank Test: p= 0.43, t-test: p=.61) .

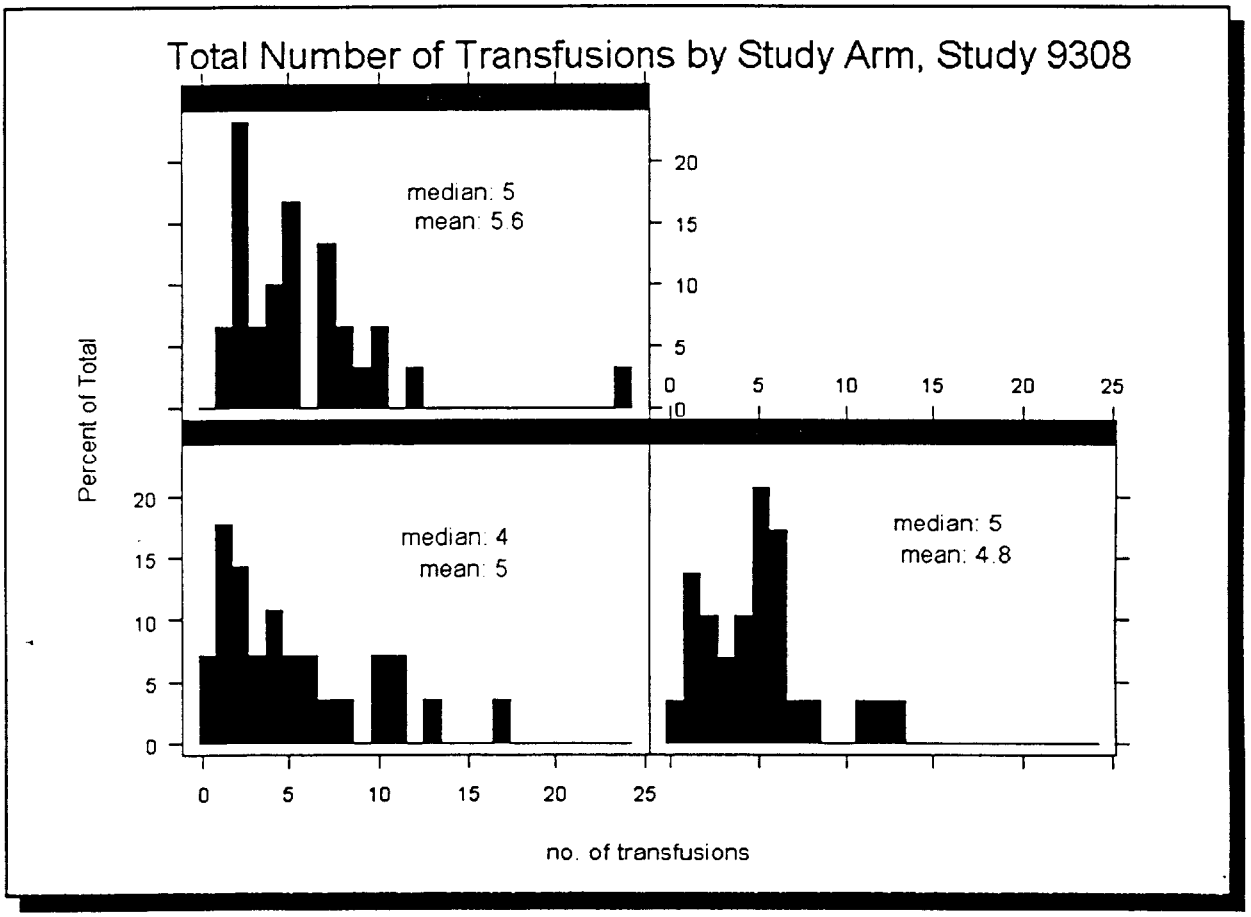


Figure 1: total number of transfusions, based on 87 patients who received study drug and had follow-up

When these summaries and analyses were prepared, the expectation was that the patients were balanced between arms with respect to the number of transfusions they would have received. However, an essential feature of this study is that patients were being compared with their prior chemotherapy history, in particular, with their previous cycle. Any treatment effect is best measured by looking at each patient's differences from the previous cycle. The electronic line listings provided numbers of platelet transfusions and red blood cell transfusions following the previous cycle. Since the patients were not on study during this cycle, it is probable that the transfusions given were not consistent with the protocol. The statistical reviewer did confirm that all but two patients (1 and 83) had received platelet transfusions with platelet counts below 20,000/μL; however, it is not known which platelet counts may have triggered subsequent transfusions.

Since this was purely an exploratory analysis to be used as supportive to the primary findings, comparisons were only made between placebo and the high dose arm. The summary statistics are shown in the tables below:

Table 7: Per-Patient Increase in Platelet Transfusions from Cycle X to Cycle X+1, Study 9308

	Placebo (N=30)	**High Dose rhIL-11 (N=28)**
median (Q1-Q3)	1 (0, 1.75)	0 (-1, 1.5)
mean	0.67	0.75
range	(-14, 6)	(-4, 11)

Table 8: Per-Patient Increase in Red Blood Cell Transfusions from Cycle X to Cycle X+1, Study 9308

	Placebo (N=30)	**High Dose rhIL-11 (N=28)**
median (Q1-Q3)	0 (-1, 1)	0 (-1, 1.25)
mean	0	0.25
range	(-3,3)	(-4, 3)

Two-sided Wilcoxon tests were used to compare the two randomized arms with respect to differences in the number of transfusions from Cycle X to Cycle X+1. The p-values were 0.11 and 0.63 for the change in platelet transfusions and red blood cell transfusions, respectively.

VI. Safety Analyses

A. Serious Adverse Events

Fifteen patients had serious adverse events in the course of the trial. Narrative summaries of these patients is provided in volume 48. From the table below, one sees a dose-response trend, with the high dose associated with the highest number of serious adverse events. A Cochran-Armitage trend test was performed on the data below using StatXact. The 2-sided exact p-value was .02. Although this was a retrospective analysis, the data are suggestive enough that toxicities in future studies should be closely monitored.

Table 9: Numbers of Patients experiencing Serious Adverse Events, Study 9308

Dose Group	Experienced a serious AE	No serious AE
Placebo	2	28
low dose	4	27
high dose	9	23

B. Time to Platelet Recovery and Time to ANC Recovery

The statistical reviewer confirmed the sponsor's analysis that the administration of rhIL-11 is not associated with an increased time to platelet recovery nor an increased time to neutrophil recovery. Time to platelet recovery was defined in the FDA analysis as the first day of a sustained platelet count above 20,000 /μL unassociated with a platelet transfusion. Platelet counts for each patient were plotted and day to platelet recovery of 20,000/μL was determined in a blinded review by the clinical reviewer. Time to ANC recovery was defined as the first day of a sustained ANC of 500/μL. These times were also computed for each patient by the FDA review team. It was noted that the FDA determinations were slightly different than the sponsor's determinations. A summary of the data and a time-to-event analysis are presented in the clinical review.

STUDY 9416

I. BACKGROUND

Study 9416 was a multi-center double-blind placebo controlled trial designed to evaluate the ability of rhIL-11 to prevent severe thrombocytopenia in patients with breast cancer undergoing mylosuppressive chemotherapy. The 77 enrolled patients at a total of 14 study sites were randomized to either placebo or rhIL-11; the randomization was stratified by prior/no prior

chemotherapy and by site. Each patient was scheduled to undergo at least 2 cycles of chemotherapy. The day after receiving the chemotherapy, patients began on study drug or placebo and were followed for at least 30 days. After two cycles, patients could go on an open-label extension of the study and receive drug following additional chemotherapy cycles. The primary endpoint in this study was the incidence of platelet transfusions over 2 cycles, which were given when platelet counts were ≤20,000/μL.

II. Patient Disposition

Of the 77 enrolled patients, 64 patients continued to the second cycle of chemotherapy. Of the 13 patients who did not continue, 5 were transfused during the first cycle. There were, therefore, 8 patients who received no transfusions in the first cycle, but dropped out of the study before the second cycle of chemotherapy. Six (6) of these patients had been randomized to receive placebo, while two (2) had received the study drug. Of the 64 patients who went on the second cycle, 2 patients were not considered by the sponsor to be evaluable: patient 33 had 3 consecutive missing platelet counts on days 18-21 when the count on day 17 was below 50,000/μL, and patient 86 had a platelet count of 20,000/μL on day 13 without an associated platelet transfusion.

III. Sponsor's Analysis

The sponsor based the primary endpoint analysis on three populations. The principal analysis was the evaluable patient population, consisting of the 67 patients: 62 patients began the second cycle of chemotherapy, had no protocol violation, and had platelet counts measured over both cycles; and 5 patients who were transfused in the first cycle, but did not receive a second cycle of chemotherapy. Supporting analysis were done on the intent-to-treat population, consisting of all 77 enrolled patients, and on the "completers" population consisting of the 62 patients who underwent two cycles of chemotherapy.

This reviewer verified from the electronic line listings that, in the evaluable patient population, that the number of patients in each arm listed as having received a platelet transfusion agreed with the summary statistics in the study report. It was also confirmed that 13 patients did not continue to the second cycle, and five (5) of those received a platelet transfusion following the first cycle.

It was not stated prospectively in the amended protocol (Volume 57, Appendix A) in the intent-to-treat analysis if patients not completing the study would be considered as treatment successes or treatment failures. Because of the clear imbalance between the two arms (6 in placebo vs. 2 in rhIL-11), the decision to place all of these patients as either treatment successes or treatment failure would have a sizeable impact on the p-value. It is usual that the intent-to-treat analysis represents a more conservative estimate of treatment effect. However, the sponsor chose to categorize the eight (8) patients as treatment failures, thereby casting the study drug in the best possible light. The two patients with protocol violations were considered as successes for this analysis. In the study report, it is stated (Volume 57, section 8.7.2) that these decisions were

made before the treatment assignments were unmasked.

A summary of the sponsor's analyses for the primary endpoint appears in the table below. The p-value is based upon a two-sided Fisher's Exact Test.

Table 10: Sponsor's Analyses of Primary Endpoint, Overall Incidence of Transfusion, Study 9416

	Placebo Arm (successes/total)	**rhIL-11 Arm (successes/total)**	**p-value (Fisher's Exact Test)**
Evaluable Patient Population	14/30	26/37	0.08
Intent-to-Treat Population	15/37	27/40	0.02
Completing Patient Population	14/29	26/33	0.02

IV. Reviewer's Analysis of Primary Endpoint

A. Comments on Evaluable Patient Analysis:

(i) The medical reviewer noted that the two patients (# 33 and # 86) who were excluded from the evaluable patient analysis should have been included for the following reasons:

Upon examining daily platelet counts, Patient 33 (Placebo) had counts on days 8, 10, 12, 15, 17 and 22. The lowest count was 34,000 /μL on day 17. Other patients were not excluded, even though there were similar gaps in platelet counts. For example, in cycle 2, patient 90 (Placebo) had a platelet count of 35,000/μL on day 11 and 24,000/μL on day 15, and no counts in between. In cycle 1, patient 89 (rhIL-11) had counts only on days 1, 3, 8 and 11. Patient 129 (rhIL-11) had counts only on days 1, 4 and 6. Although patient 129 was not considered to be evaluable in the primary analysis because she did not continue into cycle 2, she was considered to be evaluable in secondary analysis in which only the cycle 1 outcomes were compared. Patient 127 (rhIL-11) had a count of 22,000/μL on day 12 of the first cycle, but had a missing count on day 13. We will clarify with the sponsor why patient 33 was not considered to be evaluable, but the others were. The medical reviewer was comfortable considering this patient to be evaluable and a treatment success.

Patient 86 (rhIL-11) was considered to be unevaluable because she had a platelet count of 20,000/μL on day 13 with no associated transfusion. The intended drug effect that the study is measuring is the reduction in the incidence of severe thrombocytopenia or bleeding, and a platelet transfusion is a measure of that incidence. Although the Agency acknowledges that there was a major protocol violation in this case, it is also true that this patient had severe

thrombocytopenia following cycle 2 and therefore that patient's endpoint was reached. The medical reviewer considered this patient to be evaluable and a treatment failure.

(ii) When data are missing, and it is not understood why they are missing, it is appropriate to perform several analyses to test the robustness of the conclusions. The analysis based upon the evaluable patient population is valid under the assumption that the patients with complete data are representative of enrolled patient population. This assumption implies that the responses of the patients with incomplete data would have followed the same pattern of responses as that seen in the patients with complete data. When this assumption holds, inference based upon this subset can be extended to the enrolled population. This sort of "missingness" is called "missing at random" or "missing completely at random". When data are not missing at random, that is, when responses are missing because of what the responses would have been, then inferences based upon an evaluable subset may not be valid.

In an evaluable patient analysis, the proportion of transfused patients who underwent two cycles of chemotherapy is an unbiased estimate of the overall transfusion rate. Including the five patients who were transfused but dropped out after the first cycle overestimates the overall transfusion rate, because it selectively excludes the patients who were not transfused but did not continue. In this study, there were 13 patients (6 in the rhIL-11 arm and 7 in the placebo arm) who did not undergo a second cycle of chemotherapy. Four of the six patients randomized to rhIL-11 were transfused, whereas only 1 of the 7 patients randomized to placebo were transfused. Although the estimate of the overall transfusion rate computed by omitting the thirteen patients is unbiased, the extreme imbalance in the arms of the incidence of transfusion among the dropouts leads to poor estimates for the transfusion rates in each of the randomized arms. In particular, the transfusion rate in the rhIL-11 arm is underestimated when the 5 transfused patients are not included in the estimate, and the transfusion rate in the placebo arm is overestimated when the eight patients who were not transfused are not included. A revised evaluable patient analysis was therefore performed by the FDA using 64 patients with complete data and the 13 patients with partial data.

There are several approaches to doing an evaluable patient analysis when data are missing. We consider two approaches here, which lead to similar conclusions. In both approaches, unbiased estimates for the proportion of patients transfused are computed, along with estimates of the standard deviations. The two proportions are subsequently compared under the assumption that the estimated proportions are normally distributed.

The first approach is based upon estimation using the maximum likelihood principle. The principal reference is "Statistical Analysis with Missing Data", by R. Little and D. Rubin, page 173, Example 9.1. Associated with each arm is a 2x2 contingency table, with rows being the outcome following the first chemotherapy cycle and columns representing the outcome of the second chemotherapy cycle. These data are summarized in the two tables below:

Table 11: Incidence of Transfusion, rhIL-11 arm (N=40), Study 9416

	avoided transfusion/ Cycle 2	**transfused/ Cycle 2**	**did not continue**
avoided transfusion/ Cycle 1	26	4	2
transfused/ Cycle 1	0	4	4

Table 12: Incidence of Transfusion, Placebo arm (N=37), Study 9416

	avoided transfusion/ Cycle 2	**transfused/ Cycle 2**	**did not continue**
avoided transfusion/ Cycle 1	15	6	6
transfused/ Cycle 1	1	8	1

The third column contains the 13 patients with incomplete data. The goal is to find ML estimates for the probabilities of being in each of the cells in the first two columns. The estimated probability of the first row, first column (cell (1,1)) will give estimates for the transfusion avoidance rates over the two cycles in each of the arms. Let c_{ij} (I=1,2, j=1,2) be the number of subjects in the ij^{th} cell, and let r_i be the number of subjects in the i^{th} row with partial data, i.e. in the third column. The ML estimate of the probability of being in cell (1,1) is given by:

$$p_{11} = \frac{c_{11} + \frac{c_{11}}{c_{11}+c_{12}} r_1}{N}$$

so that the missing data is distributed to the cells in proportion to the expected number given the complete data. The estimated variance for the estimator p_{11} is given on page 180 of the above cited reference. As might be expected, the estimated variance is smaller than $p_{11}(1-p_{11})/m$, where m is the total number of complete patients, but larger than $p_{11}(1-p_{11})/N$, where N is the total number of enrolled patients. For the placebo arm, the estimated transfusion avoidance rate (p_{11}) was 0.52 with an estimated variance of 0.0077. For the rhIL-11 arm, the estimated transfusion avoidance rate was 0.69 with an estimated variance of 0.0065. The Z-statistic, as noted below,

$$\frac{0.69-0.52}{\sqrt{0.0065+0.0077}} = 1.4$$

corresponds to a two-sided p-value of 0.16.

Another approach to the evaluable patient analysis utilizes a survival analysis model. In this case, a "survivor" is a subject who avoided transfusions. Although the usual survival analysis compares time to transfusion, this analysis compares survival estimates following two cycles of chemotherapy. The variance estimates for the estimated survival are based upon Greenwood's formula. This type of approach fits nicely with the structure of the data, since one has censored observations following the first chemotherapy cycle, as well as transfusion events in the first cycle with no follow-up in the second cycle. The transfusion avoidance estimate in the two arms are simply the Kaplan-Meier estimates of survival, and are computed below:

For the placebo arm: $p_{placebo}= 27/37 * 15/21 = 0.52$, Var $(p_{placebo})= 0.0079$

For the rhIL-11 arm: $p_{rhIL-11}= 32/40* 26/30 = 0.69$, Var $(p_{rhIL-11})= 0.0054$

The normalized Z-statistic comparing the two proportions,

$$\frac{0.69-0.52}{\sqrt{0.0054+0.0079}} = 1.47$$

corresponds to a two-sided p-value of 0.14.

It should be noted that both estimates of the proportion responders are ML estimates of the same likelihood function, and are therefore equivalent.

B. Comments on Intent-to-Treat Analysis: As stated above, the purpose of an intent-to-treat analysis is to reassess a possible treatment effect in a more conservative evaluation. If patients drop out a study because of drug toxicities or inconveniences associated with the drug, one should include these patients in order to capture this information. Often, it is not clear whether the reason the patient dropped out was related to the study drug, but when there is a differential drop out rate, one should give more credence to the intent-to-treat analysis.

In this case, there were 7 patients in the Placebo Arm and 6 patients in the rhIL-11 arm who dropped out after the first chemotherapy cycle. However, 6/7 placebo patients had no transfusion, whereas only 2/6 patients in the rhIL-11 arm avoided transfusion. That is, of the 8 patients with no transfusions following the first chemotherapy cycle, 6 (patient #'s 20, 27, 50, 91, 122, 132) were in the placebo arm and 2 (patient #'s 25, 129) were in the rhIL-11 arm. The medical reviewer has reviewed each patient's case report form carefully and has confirmed the various

reasons for discontinuation.

The most conservative analysis would assign the 2 patients in the rhIL-11 the worst possible outcome, that is, treatment failure, while assigning the 6 patient in the placebo arm the best possible outcome, that is, treatment success. One would hope that the conclusion or the suggestion of a treatment effect would be robust to this most ungenerous way of imputing outcomes in patients with unknown outcome. Using these assignments, the imputed success rate of 21/37 (57%) in the placebo arm was similar to the imputed success rate of 26/40 (65%) in the rhIL-11 arm, and not statistically significant.

Another conservative analysis would categorize all patients with uncertain outcomes as successes. This is in contrast to the sponsor's analysis of assigning these patients as failures and reporting a more favorable p-value than in the evaluable patient analysis. By this count 21/37 (57%) of placebo patients were counted as treatment successes, while 28/40 (70%) of rhIL-11 patients counted among the treatment successes. Although a higher percent of patients on study drug avoided transfusion, this difference was not statistically significant by the two-sided Fisher's Exact Test (p=0.25). This p-value contrasts sharply with the p-value generated by the sponsor in their ITT analysis (p=0.02). Such large differences, which would naturally lead to different conclusions must be carefully examined.

A summary of the differences in the overall incidence of transfusion in the two arms using the FDA patient evaluations is presented in the table below:

Table 13: FDA's Analyses of Primary Endpoint, Overall Incidence of Transfusion, Study 9416

	Placebo Arm (successes/total)	**rhIL-11 Arm (successes/total)**	**p-value (Fisher's Exact Test)**
Evaluable Patient Population	15/31	26/38	0.14
Intent-to-Treat (ITT) Population	21/37	28/40	0.25
ITT Population (worst case scenario)	21/37	26/40	0.49

V. Secondary Analysis: Accounting for Stratification Variables

A. Site Differences: In this study, patients were randomized by site, and consequently, treatment groups were well-balanced within centers. Unlike the design of study 9308, differences seen in overall transfusion rates among treatment groups could not be attributed to center differences.

However, it is still worth examining the treatment effect center by center, to affirm that the treatment effect is consistent, or homogenous, between centers. If there is a lack of homogeneity, that is, if there are centers in which the trends favoring efficacy go in the reverse direction from other centers, then it behooves the sponsor and the FDA to understand why the drug appears to act favorably in one set of circumstances, but unfavorably in another, and whether these differences in the measure of treatment effect are important enough to question the overall result of the trial. On the other hand, when the treatment effect appears to be relatively homogenous among centers, then either an analysis based upon the overall incidence of transfusion, or a stratified analysis by center should give similar p-values.

This reviewer investigated the possibility that a treatment effect was not homogenous across centers. There were 14 investigative sites, but patients were not uniformly distributed across sites. The two largest centers (inv. site #s 108 and 111) had 13 patients each, whereas 5 of the centers (inv. site #s 109, 112, 126, 135, 151) had fewer than 3 patients. Of these 5 centers, only site # 112 had a patient in both treatment groups.

In general, however, the numbers of patients randomized to each treatment group were well-balanced in each of the centers. So although the incidence of transfusion varied widely from center to center (from 0% in site # 125 with 9 evaluable patients to 83% at site #38 with 6 evaluable patients), any difference seen between the incidence of transfusion in the two treatment groups cannot be attributed to site variability. A breakdown of the incidence of transfusion by site is summarized in the table below:

Table 14: Incidence of Transfusion by Site, Study 9416

	transfused	**avoided transfusion**	**not evaluable**
Site # 108 (N=13)			
Placebo	1	5	1
rhIL-11	0	6	0
Site # 111 (N=13)			
Placebo	3	2	1
rhIL-11	3	3	1
Site #125 (N=9)			
Placebo	0	4	0
rhIL-11	0	5	0
Site #38 (N=8)			
Placebo	3	0	1
rhIL-11	2	1	1

	transfused	avoided transfusion	not evaluable
Site #110 (N=8)			
Placebo	2	1	1
rhIL-11	0	4	0
Site #30 (N=5)			
Placebo	2	0	0
rhIL-11	2	1	0
Site #107 (N=5)			
Placebo	1	1	1
rhIL-11	0	2	0
Site #152 (N=5)			
Placebo	2	0	1
rhIL-11	1	1	0
Site #75 (N=3)			
Placebo	1	0	0
rhIL-11	0	2	0
Site #112 (N=2)			
Placebo	1	0	0
rhIL-11	1	0	0
Site #135 (N=2)			
Placebo	0	2	0
rhIL-11	0	0	0
Site #151 (N=2)			
Placebo	0	0	0
rhIL-11	2	0	0
Site #109 (N=1)			
Placebo	0	0	0
rhIL-11	1	0	0
Site #126 (N=1)			
Placebo	0	0	0
rhIL-11	0	1	0

Of the sites with at least 5 enrolled patients, only site #110 shows any appreciable treatment effect. The effect in every other center can be described as marginal, at best. However, there was

no center at which the placebo patients fared better with respect to the primary endpoint. A test of homogeneity of treatment effect across centers was performed using StatXact. A p-value of .77 does not support the lack of homogeneity across centers. It must be noted, however, that this analysis excludes sites 109, 112, 125, 126, 135, and 151 comprising 17 patients, because there are zeros in the margins of these tables.

The analysis above suggests that the treatment effect is relatively homogenous across centers, and therefore either an analysis comparing overall rates among arms or an stratified analysis adjusting for center effects is appropriate. The results of the analysis comparing overall rates is shown on Table 2, and had a p-value of 0.14. The stratified analysis was not done, because this analysis would have excluded the 17 patients from the sites with zeros in the margins. It was felt that the results of these 17 patients should be part of any efficacy analysis.

B. Homogeneity of Treatment Effect by Prior Chemotherapy: Since patients were stratified by prior/no prior chemotherapy, one can also ask where the treatment effect seen was similar in these two groups. The two tables below show a breakdown of the subjects by prior chemotherapy. Although the numbers are very small in the prior chemotherapy subgroup, the differences in the incidence of transfusion between the rhIL-11 arm and the placebo arm are quite striking (38% vs. 70%) although not significant using Fisher's Exact Test (p=.21). In contrast, there is less difference seen in the proportion transfused between the two treatment groups in the no prior chemotherapy subpopulation (26% (7/27) rhIL-11 vs. 33% (9/27) placebo). Although this subpopulation represents 54 (70%) of the 77 enrolled patients, a clear treatment effect was not evident (Fisher's Exact Test, p=.77). Although there appears to be a treatment interaction (meaning that a treatment effect is seen in one subpopulation but not the other), an exact test for treatment interaction (using Zelen's statistic, StatXact) was not statistically significant (p=.27). In the opinion of this reviewer, however, the hypothesis that rhIL-11 may have greater efficacy in subjects with prior chemotherapy remains viable and should be re-examined using larger sample sizes.

Table 15: Incidence of Transfusion in Evaluable Patients with No Prior Chemotherapy, Study 9416

	transfused	**avoided transfusion**	**not evaluable**	**total**
Placebo	9	14	4	27
rhIL-11	7	19	1	27
total	16	33	5	54

Table 16: Incidence of Transfusion in Evaluable Patients with Prior Chemotherapy, Study 9416

	transfused	avoided transfusion	not evaluable	total
Placebo	7	1	2	10
rhIL-11	5	7	1	13
total	12	8	3	23

C. Numbers of Platelet Transfusions and Red Blood Cell Transfusions

Unlike Study 9308, missed transfusions were not documented, so the total number of transfusions required was probably somewhat underestimated. It should be noted that although the mean number of transfusions drops in the rhIL-11 arm from cycle 1 to cycle 2, only one patient (90) experienced a reduction in number of transfusions from 1 to 0, although given that this patient missed several platelet counts, it cannot be determined if a platelet transfusion was required. The decrease in the mean is only a reflection of the fact that 4 out of the 8 patients transfused in that arm dropped out after the first cycle. Although cycle 2 summary data are reported here, the estimates in the rhIL-11 arm may be underestimated, and estimates in the placebo arm may be overestimated.

Table 17: Number of Platelet Transfusions, Study 9416

Study Arm	Cycle 1				Cycle 2			
	median	mean	range	# evaluated	median	mean	range	# evaluated
Placebo	0	0.43	0-6	37	0	1.5	0-11	30
rhIL-11	0	0.33	0-4	40	0	0.29	0-3	34

Table 18: Number of Red Blood Cell Transfusions, Study 9416

Study Arm	Cycle 1				Cycle 2			
	median	mean	range	# evaluated	median	mean	range	# evaluated
Placebo	0	0.43	0-3	37	0	0.6	0-17	30
rhIL-11	0	0.5	0-3	40	1	0.74	0-6	34

SAFETY ANALYSES

All of the following analyses were done for each of the two cycles. However, it should be noted that the patients who dropped out after cycle 1 were unbalanced in the two arms with respect to transfusion requirements. Of the 8 rhIL-11 patients who required a transfusion in the first cycle, four did not continue into the second cycle. Moreover, the remaining four patients were transfused in the second cycle. Only two of the rhIL-11 patients who dropped out after the first cycle did not require a transfusion. If requiring a transfusion is an indicator of being sicker at baseline, then the continuing patients may not be representative of the patients originally randomized to this group. On the other hand, 6 of the 27 patients who avoided transfusion in the placebo arm did not continue into cycle 2. Of the 10 placebo patients who were transfused in the first cycle, only one did not go on to the second cycle. In contrast to the rhIL-11 arm, the patients in the placebo arm continuing into the second cycle may have been, on average, less sick at baseline than the patients originally randomized to placebo. One should therefore be cautious about overinterpreting comparisons of the summary statistics of cycle 1 with cycle 2.

A. HEMOGLOBIN LEVELS

The clinical reviewer expressed a concern that hemoglobin levels may be lower in the rhIL-11 arm. For each cycle, all hemoglobin measurements were plotted and a smoothing spline was fit for each of the arms. The graphs are presented in the appendix. Mean concentrations for a subset of the study days are displayed in the below tables.

Table 18: Mean Hgb Concentrations: Cycle 1, Study 9416

Study Arm	Mean Hgb Concentration (g/dL) on Study Day (Cycle 1)						
	Day 0	Day 2	Day 4	Day 7	Day 14	Day 20	Day 26
Placebo	11.7 (n= 28)	11.2 (n= 23)	11.7 (n= 20)	10.4 (n= 15)	9.8 (n= 20)	10.3 (n= 8)	NA (n= 0)
rhIL-11	11.6 (n= 35)	10.9 (n= 22)	10.8 (n= 17)	9.5 (n= 22)	9.8 (n= 21)	9.8 (n= 10)	10.1 (n= 2)

Table 19: Mean Hgb Concentrations: Cycle 2, Study 9416

Study Arm	Mean Hgb Concentration (g/dL) on Study Day (Cycle 2)						
	Day 0	Day 2	Day 4	Day 7	Day 14	Day 20	Day 26
Placebo	10.6 (n=27)	10.3 (n=15)	10.5 (n=16)	9.9 (n=19)	9.3 (n=18)	10.0 (n=8)	9.5 (n=2)
rhIL-11	10.5 (n=27)	9.8 (n=17)	9.9 (n=13)	9.1 (n=21)	9.7 (n=23)	9.8 (n=7)	10.2 (n=1)

B. DAYS TO ANC RECOVERY

The statistical reviewer confirmed the sponsor's conclusions that rhIL-11 had no adverse effect on ANC recovery. Time to ANC of 500/μL was computed for each patient as the number of days from the first day of chemotherapy until the first day of a sustained ANC over 500/μL using the ANC data in the electronic line listings. For approximately 15% of the subjects, there were missing ANC data during critical days and no count below 500/μL. For these patients, total white counts were examined. Any count below 1,000/ was counted as a day of severe neutropenia. In the sponsor's determination of time to ANC recovery and duration of severe neutropenia, missing values were ignored, and those patients with no counts below 500 /μL were listed as having experienced no severe neutropenia. Moreover, time to ANC recovery was measured by the sponsor started from the first day of drug administration. Given these discrepancies, the FDA summary data differed somewhat from the sponsor's. Nevertheless, there was no statistically relevant difference seen between the two arms in either cycle.

Table 20: Days to ANC Recovery, Study 9416

Study Arm	Cycle 1				Cycle 2			
	median (days)	mean (days)	range (days)	# unevaluable	median (days)	mean (days)	range (days)	# unevaluable
Placebo	13	13.0	11-20	0	12.5	12.6	0-17	7
rhIL-11	12	12.3	0-16	1	12	12.4	11-15	6

STUDY 9313

I. BACKGROUND

Study 9313 was a single-center randomized double-blind placebo-controlled three arm phase 2 study in women with breast cancer undergoing high-dose chemotherapy with autologous bone marrow transplantation. A total of 80 women were randomized to receive either placebo, low dose rhIL-11 (25 mcg/kg) or high dose rhIL-11 (50 mcg/kg). Seventy-five patients are included in the safety analysis, since they received at least one dose of study drug.

II. SAFETY ANALYSIS

The review of this study was limited to a comparison of the days to hematopoietic recovery among the three arms. Time to ANC recovery and time to platelet recovery were measured as

days from the start of study drug until the first day of recovery: for ANC, a sustained count ≥500/μL, and for platelets, a sustained of >20,000/μL unassociated with a platelet transfusion. Although summary statistics were available from the sponsor, these numbers were determined for each patient by the clinical reviewer and the statistician from the electronic line listings. The summary of these data is shown below:

Table 22: Days to ANC Recovery, Study 9313

	Placebo (N=25)	**Low Dose rhIL-11 (N=26)**	**High Dose rhIL-11 (N=24)**
median	11 days	11 days	11 days
range	10-31 days	10-31 days	9-24 days
# patients not recovering	1	0	1

Table 23: Days to Platelet Recovery, Study 9313

	Placebo (N=25)	**Low Dose rhIL-11 (N=26)**	**High Dose rhIL-11 (N=24)**
median	15 days	14 days	17 days
range	12-83 days	11-47 days	11-48 days
# patients not recovering	3	1	0

Since the high dose of rhIL-11 is the proposed dose for the indication, only the high dose and the placebo were compared with respect to time to platelet and neutrophil recovery. A log-rank test was used. The test of no difference in time to ANC recovery yielded p-value of 0.69. In addition, the p-value associated with no difference in time to platelet recovery was 0.90; neither analysis was suggestive of a difference in time to hematopoietic recovery between the two arms.

In addition, the number of platelet transfusions and units of red blood cell transfusions in each arm were tabulated and compared. Although means and standard deviations were reported in the sponsor's analysis, it should be noted that the data are highly skewed in the placebo arm, and the mean and standard deviation fail to adequately summarize the data. Other summary statistics are included in the table below:

Table 24: Number of Platelet Transfusions by Study Arm, Study 9313

	Placebo (N=25)	Low Dose rhIL-11 (N=26)	High Dose rhIL-11 (N=24)
median (Q1,Q3)	8 (6,12)	7 (6,9)	9 (6, 11)
mean (s.d.)	11.4 (10.2)	8.0 (4.6)	8.8 (3.5)
range	3-46	3-22	4-17

Table 25: Number of Units of Red Blood Cell Transfusions, Study 9313

	Placebo (N=25)	Low Dose rhIL-11 (N=26)	High Dose rhIL-11 (N=24)
median (Q1, Q3)	3 (2,6)	4 (3,6)	5 (4,7)
mean (s.d.)	4.5 (4.0)	4.2 (2.2)	5.0 (2.2)
range	0-17	0-8	0-9

Differences between high dose arm and the placebo arm were tested using the two-sided Wilcoxon Rank Sum Test. The p-values were 0.96 and 0.13 for the comparison of numbers of platelet transfusions and numbers of units of red blood cell transfusions, respectively. Neither test provided evidence of a difference in the number of transfusions required between the two arms.

CONCLUSIONS

Study 9313: Although Study 9313 was not submitted to support the efficacy of rhIL-11, it provided important additional evidence of the limitations of rhIL-11 in preventing severe thrombocytopenia. It was consistent with the other two supportive studies in that there was no evidence that the drug had an adverse effect on time to neutrophil recovery or time to platelet recovery.

Study 9416: This study was more straightforward than 9308, because all patients received the same chemotherapy. Few adverse events were seen in this study, and all patients were evaluable over at least one cycle. Although there was not a statistically significant difference in the two arms in the primary endpoint, the results of this study were consistent with study 9308, and support the activity of rhIL-11 in reducing the need for platelet transfusions.

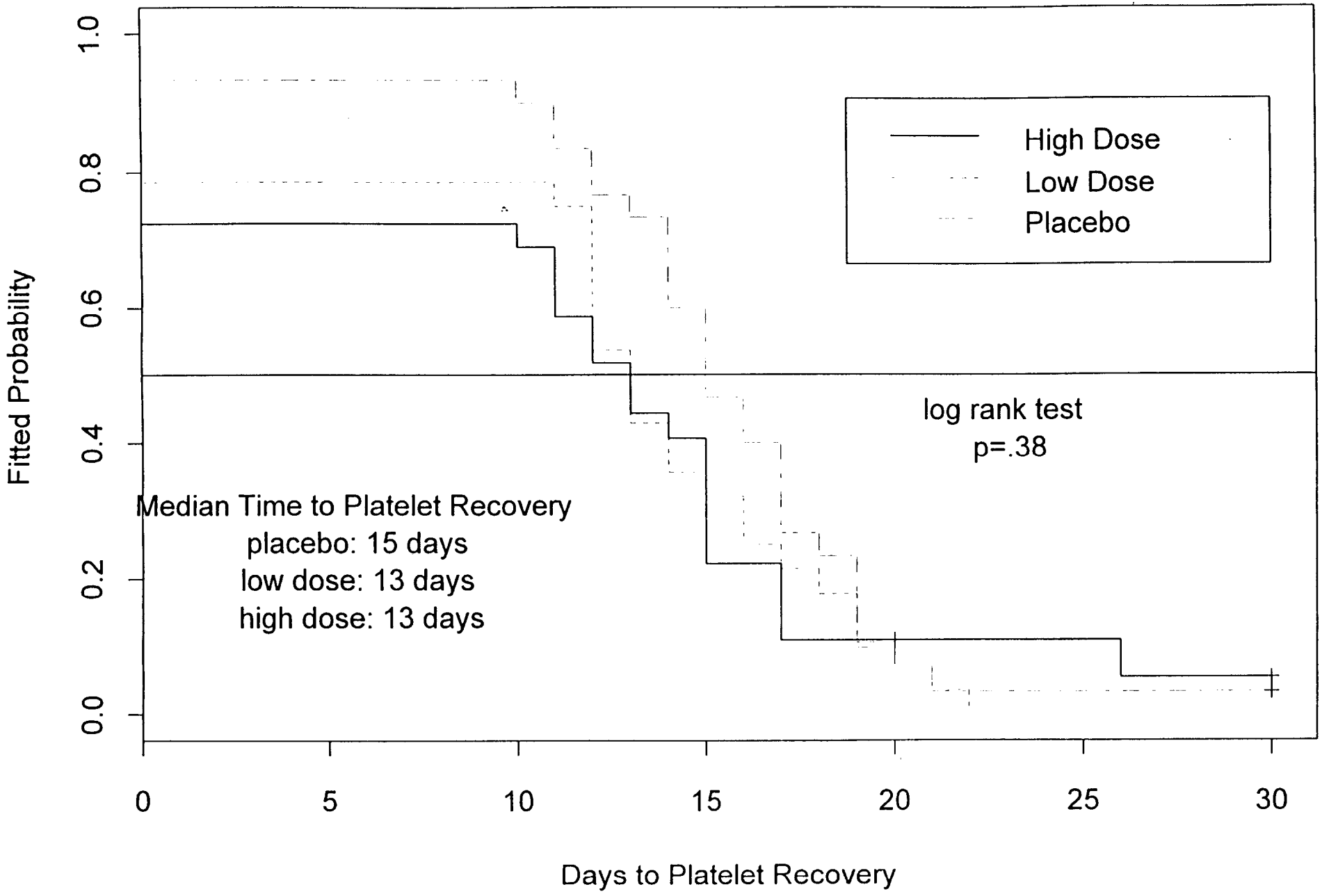
Study 9308: Study 9308 is the one study which statistically differentiated the patients in the placebo arm from the patients in the rhIL-11 arm. However the number of serious adverse events seen was also notably higher in the high dose rhIL-11 arm. The adverse events associated with this drug include edema and atrial fibrillation. If the magnitude of these events could be

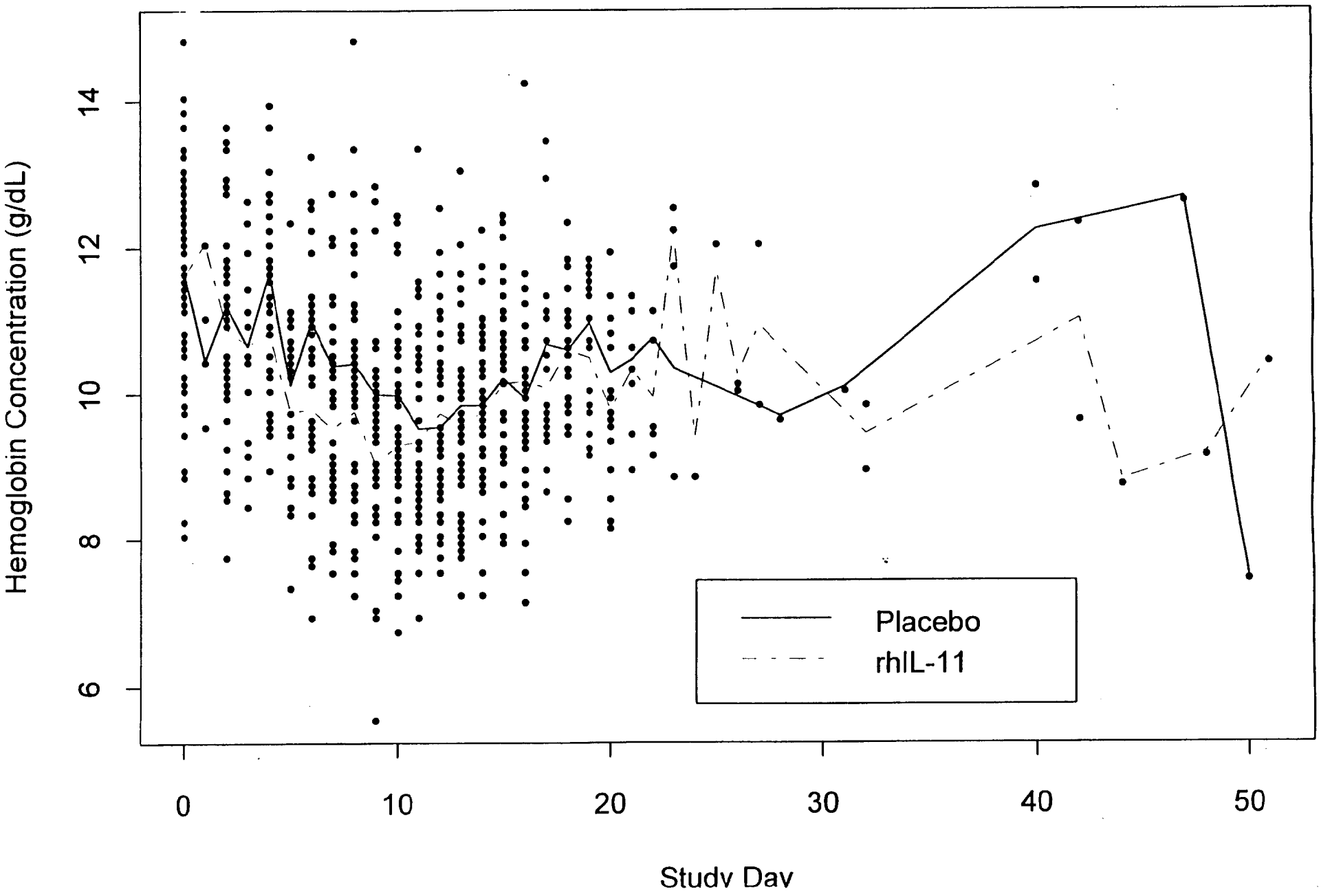
quantified and contrasted with the benefits of the drug, then a risk-benefit assessment could be made. In considering the risk-benefit ratio, one should also consider that the potential benefit lies in the ability to re-treat these patients with the same large doses of chemotherapy. Severe thrombocytopenia can also be reduced by reducing the dose of chemotherapy, and the relative benefits of being able to maintain the same dose of chemotherapy is not well understood. If a reduced dose of chemotherapy is as effective in preventing disease progression as maintaining the chemotherapy dose, and is rather more toxic, then rhIL-11 has no important role as an adjuvant in this setting.

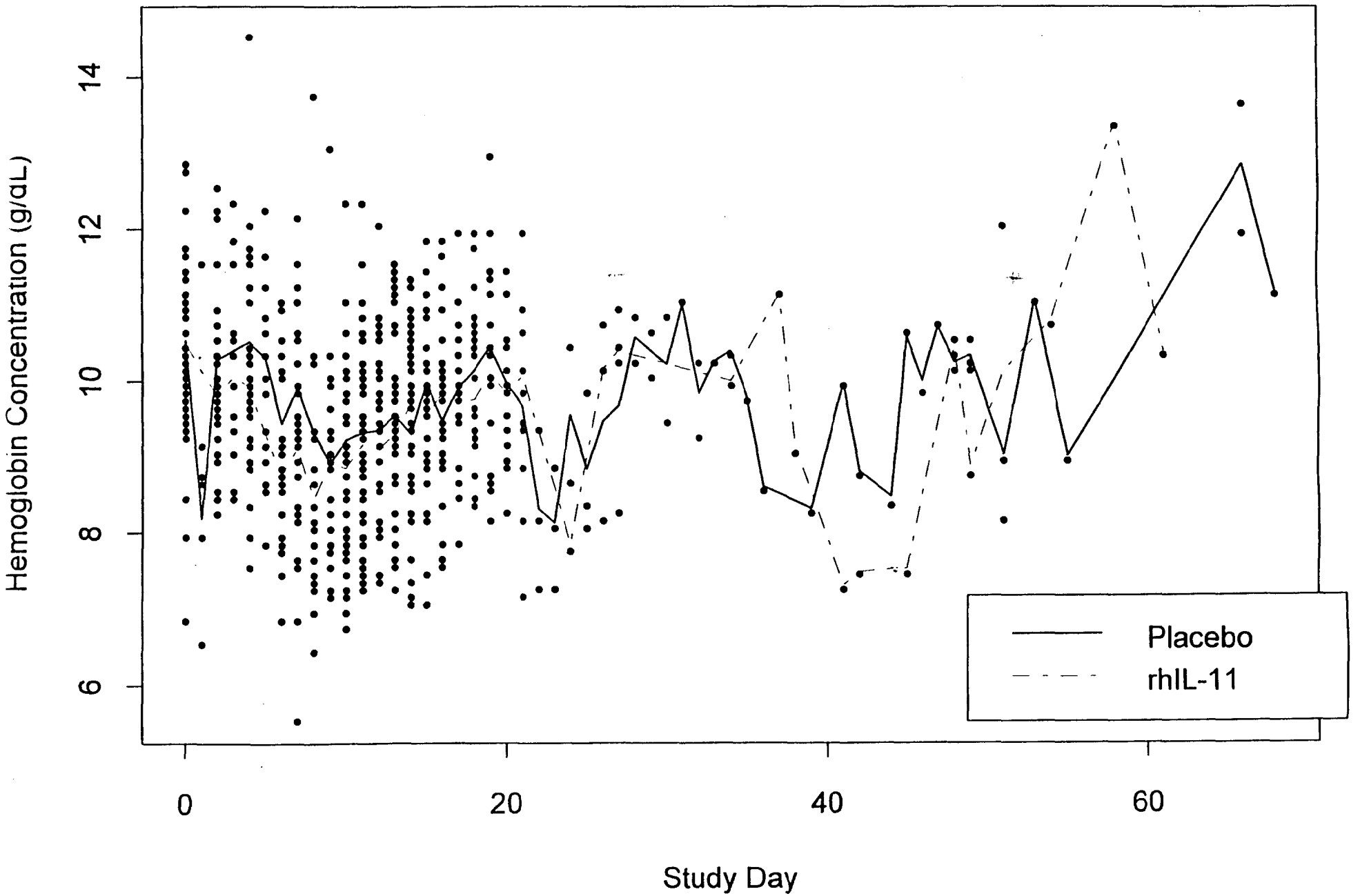
Although rhIL-11 is currently the only drug shown to be effective at preventing severe thrombocytopenia, platelet transfusions can and have been used in the treatment and prophylaxis of severe thrombocytopenia. If there is consensus that severe thrombocytopenia should be treated, then the risk-benefit assessment should also consider both the associated risks and the availability of platelet transfusions. Suppose that platelet transfusions pose fewer potential risks than rhIL-11 and are readily available. It could then be argued that rhIL-11 is not a safe enough product for licensure. If, on the other hand, if platelet transfusions are not an available option for a patient, then rhIL-11 may be safe relative to the potential risks a patient may experience if left untreated. Thus, the risk-benefit picture is, at best, a complicated one and requires careful deliberation.

cc: Chron File, HFM-210
Chris Joneckis, DARP, HFM-594
Susan Ellenberg, DBE, HFM-210
Peter A. Lachenbruch, DBE, HFM-215

Kaplan-Meier Estimates for Time to Platelet Recovery
Study 9308, Genetics Institute




Hgb Concentration by Study Day (g/dL): Cycle 1
Study 9416, Genetics Institute
Hemoglobin Concentration (g/dL)
6
8
10
12
14
0
10
20
30
40
50
Study Day
Placebo
rhIL-11


Hgb Concentration by Study Day (g/dL): Cycle 2
Study 9416, Genetics Institute
Hemoglobin Concentration (g/dL)
6
8
10
12
14
0
20
40
60
Study Day
Placebo
rhIL-11